Airport Benchmarking - An Empirical Research on the Performance Measurement of German Airports with Data Envelopment Analysis

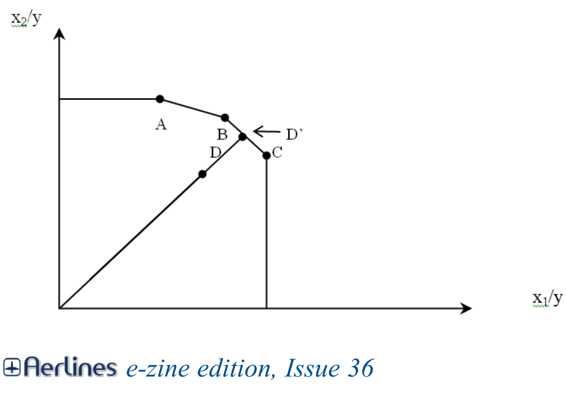
The increase in interest of benchmarking in the airport industry is not only visible on the academic side but also comes from the airport management and/or authorities who might use it as a regulatory tool. The reasons for this shift in focus are, from the managements' perspective, the increase in the number of privatizations as well as more commercialization and non-aviation related activities at airports since the deregulation of the air transport industry. This case study measures the technical efficiency of sixteen international airports in Germany from 1998-2004 with Data Envelopment Analysis (DEA) and creates a ranking of the selected airports.

By Vanessa Kamp

Based on currently available studies, German airports appear to be less financially efficient and productive compared to other airports in Europe and in particular to non-European airports. The reason for this might be the high degree of vertical integration at German airports. It is important to note though, that only a small number of international studies have included German airports; a national study on measuring the technical efficiency does not exist so far.

This case study measures the technical efficiency of sixteen international airports in Germany from 1998-2004 with Data Envelopment Analysis (DEA) and creates a ranking of the selected airports. It is one of many to measure the overall efficiency of German airports as part of the research project German Airport Performance (GAP). This paper deals with the analysis of technical and traffic data, since adequate financial data is not yet available.

Figure 1: *Data Envelopment Analysis*



Methodology and Data

Data Envelopment Analysis measures the relative efficiency according to Farrell (1957). It is a non-parametric approach that uses linear programming to construct a piece-wise linear frontier, which is determined by the efficient airports of the sample (see figure 1). The concept of measuring the technical efficiency with linear programming was first introduced by Charnes, Cooper and Rhodes (1978).

An advantage of the DEA is that it can handle multiple inputs and outputs within a single analysis without any difficulties of aggregation. Instead of weighting factor quantities as is done when measuring the Total Factor Productivity (TFP), DEA optimizes the weights with linear programming:

$$max\,\theta^{k'}$$

$$s.t.$$

$$\sum_{k=1}^{K} z^k y_m^k \geq \theta^{k'} y_m^{k'}, \mathrm{m}=1,\ldots,\mathrm{M}\ (\mathrm{Outputs})$$

$$\sum_{k=1}^{K} z^k x_n^k \leq x_n^{k'} \quad , \mathrm{n}=1,\ldots,\mathrm{N}\ (\mathrm{Inputs})$$

$$z^k \geq 0$$

where $\theta k'$ indicates the efficiency score of every airport k'. zk are the weights that are determined by the optimization process. A value of θk'=1 indicates a point on the frontier and thus a technically efficient airport. This linear programming problem must be solved K times, once for each airport, hence $\theta k'$ has to be obtained for each firm (Coelli et al 2005).

DEA can either focus on input minimization with constant outputs, or it can calculate an output maximization model by holding the inputs constant. The decision is very often up to the management to affect certain variables. Furthermore, when applying DEA, one has to assume either constant returns to scale (CRS) or variable returns to scale (VRS). This depends on whether all airports can operate at an optimal scale. If not, it is more appropriate to assume variable returns to scale, because it decomposes the technical efficiency score into a) scale inefficiency and b) 'pure' technical efficiency.

Only a single output has been used, as this case study was also part of an analysis with Stochastic Frontier Analysis (SFA), which needs the functional form of a production frontier. Here, the author decided to choose the annual passenger volume as the output for the model. Cologne-Bonn as an airport with high cargo volume has been excluded due to a lack of data. Another reason is that the work load unit (WLU) selected in a previous study did not lead to sufficient results. Even though it is the standard measure in the aviation industry, the question arises if the effort for handling a passenger is comparable to the effort of handling a 100 kg amount of cargo. Strictly speaking, fixing two outputs in certain proportions is incompatible with the optimization of the firm, namely that a firm with multiple products maximizes profits by equating the partial marginal revenues with partial marginal costs (Selten 1970).

The following inputs seemed to be the most appropriate variables after a correlation analysis and the test of significance: the number of check-in counters, the number of gates, the airport size (given in hectare), the number of runways, and the number of car parking spots.

Unfortunately, the number of employees that is available is unadjusted regarding the vertical integration, but the service level at German airports varies from airport to airport. Most airports provide labor-intensive activities, such as ground handling, which are often outsourced at airports outside of Europe. However, this service has always been provided by a third party at the Berlin airports. The different degree of vertical integration at German airports can affect an airports' relative performance, and leads to misleading results. For this reason, the number of employees has been excluded from this sample. As there is no substitutability between labor and any other factor in the production there should not be any problem with this (Pels et al 2001).

Results of Data Envelopment Analysis

For the output maximizing model, variable returns to scale were assumed due to the existence of airports with different passenger volumes. The sample includes Frankfurt with an annual number of more than 50 million passengers in 2004, and Saarbruecken, the smallest airport amongst the international airports with less than 0.5 million passengers per year.

The results indicate four of sixteen airports that operate less than 100 percent technical efficiency in the time period concerned. These are Saarbruecken and Frankfurt (see table 1), but also Stuttgart and Berlin-Tegel. There are also airports that operate efficiently almost over the whole period such as Bremen, Duesseldorf, and Muenster-Osnabrueck.

The weakest performance in the sample was identified for the airports in Hanover, Berlin-Schoenefeld and Berlin-Tempelhof with average efficiency scores of 68 percent, 45 percent and 39 percent respectively. An explanation might be the excess capacity compared to other airports with a similar throughput. Hanover as an airport with 5.1 million passengers in 2004 has the third largest airport area in Germany, thus being larger than the airport in Duesseldorf (15.1 million) , in Hamburg (9.8 million), in Stuttgart (7.4 million) and in Berlin-Tegel (11.0 million). This indicates more supplied facilities on the airside and terminal side than have to be needed to handle their passengers. Indeed, when plotting the passenger volume against the airport size, an above average input can be identified for Hanover, Leipzig, Berlin-Schoenefeld and also Munich (see Figure 2). As in Hanover, also in Berlin-Schoenefeld, the major influence of technical inefficiency may arise due to the large airport size. Until financial year 2003, this airport had a similar throughput as Bremen and Dresden of less slightly less than 2.0 million passengers but their airport size is more than twice as big as in Bremen and Dresden. However, in Berlin-Schoenefeld the technical efficiency increased from 43% to 84% in 2003/04 due to a passenger increase of 97% from 1.7 to 3.3 million. This high increase is mainly due to the higher volume of Low Cost Carrier (LCC) traffic in Berlin-Schoenefeld compared to previous years.

Considering the potential to grow of the 16 airports indicates highest rates for Hanover, Berlin-Schoenefeld, Berlin-Tempelhof but also for Leipzig in order to become more technically efficient . As with the least efficient airports, Leipzig also has a relatively large airport size which has been expanded in 2000 from 300 ha to 800 ha. For a further analysis, the DEA-model was run again but only considering 2003 and 2004. The results indicate the potential of the airports' output (here the number of passengers) to grow to become technically efficient relative to its reference sets as seen in Table 2.

Table 2: *Potential Output growth in %*

	2003			2004		
	Actual Output	Target Output	Difference	Actual Output	Target Output	Difference
BRE	1.60	1.60	0.0%	1.65	1.65	0.0%
DRS	1.50	1.70	13.33%	1.60	1.70	6.25%
DTM	1.00	1.50	50.00%	1.20	1.40	16.67%
DUS	14.20	14.40	1.41%	15.15	15.15	0.0%
FMO	1.50	1.50	0.0%	1.46	1.46	0.0%
FRA	48.11	48.11	0.0%	50.77	50.77	0.0%
HAJ	4.90	7.80	59.18%	5.20	7.90	51.92%
HAM	9.40	11.10	18.09%	9.80	11.70	19.39%
LEJ	1.90	3.30	73.68%	1.90	3.40	78.95%
MUC	24.00	27.80	15.83%	26.79	29.70	11.24%
NUE	3.20	3.40	6.25%	3.59	3.59	0.0%
SCN	0.41	0.41	0.0%	0.41	0.41	0.0%
STR	7.46	7.47	0.0%	8.70	8.70	0.0%
SXF	1.70	3.90	129.41%	3.30	3.90	18.18%
THF	0.45	1.20	166.67%	0.44	1.20	172.73%
TXL	11.06	11.06	0.0%	11.01	11.01	0.0%

Here, especially the airports in Leipzig, Berlin Schoenefeld and Berlin-Tempelhof have the highest potential to expand their output. This will certainly take place in Leipzig and Berlin-Schoenefeld.

Leipzig will expand its cargo facilities as it becomes the European hub airport for the cargo company DHL. It is therefore quite reasonable to also find an above average relationship of airport size to passenger facilities. However, the initial plan of Leipzig to build an airport for intercontinental traffic can be shown when plotting the gross terminal size against check-in counters. Leipzig has set up a huge terminal building and built a train station for long-distance trains, which is integrated in the building. Furthermore, in June 2000, they also opened an additional runway of 3,600 m in length and 60 m in width, which can be used for intercontinental flights.

The airport in Berlin-Schoenefeld was selected to become the principal Berlin airport in the future, the Berlin/Brandenburg International Airport (BBI). Compared to the two other airports in Berlin, namely Tegel and Tempelhof, it is situated more outside Berlin and is not as capacity-constrained as Tegel. The expansion of the airport is not planned to start before 2008. In 2007, the airport in Berlin-Tempelhof is planned to be closed and Berlin-Tegel will discontinue its services after the opening of BBI, which will not take place before 2011.

Another interesting point is that the airport in Hanover uses much of its airport for non-aviation activities that have nothing to do with the actual airport business, namely the handling of passengers, aircrafts and cargo. These activities are, for example, various exhibitions, and the Airport Business Park. The increasing interest in non-airport affined activities cannot be corrected with the data that has been collected so far, and thus can affect an airports' performance regarding their true capacity.

Conclusion

The DEA-results identified Frankfurt, Stuttgart and Berlin-Tegel as best practices, whereas the airports in Hanover, Berlin-Schoenefeld and Berlin Tempelhof have plenty of spare capacity.

The airport industry in Germany shows much heterogeneity and several aspects that have to be considered when measuring the technical efficiency. These are not only the staff numbers and the airport size, which were already mentioned in the text, but other considerations when costs will be included in future studies. There are, for example, the use and the costs of the terminal, as well as overcapacities on both the terminal side and the air side:

Cost allocation: the complex roof construction at the terminals of Hamburg Airport is beautiful, but will certainly not increase technical or allocative efficiency. However, the total cost of this terminal with an annual capacity of around 8 million passengers does not exceed the cost of the new terminal in Dortmund, which has an annual maximal throughput of three million passengers (Schmidt 2005). Here, the costs and the annual capacity should be set in relation to gain a meaningful ratio for the allocation of resources (see also figure 3). Other examples of quality aspects are marble floors or people movers that will also have to be considered in further analyses.

Capacity: this is another factor that needs to be investigated on both operational sides. Merely including the gross terminal size, the size of the apron or the number of runways can cause misleading results. There is, for instance, Duesseldorf Airport that can only use the capacity of one runway due to political restrictions. Hence, further analyses should also include capacity figures especially when capacity restrictions are beyond managerial control.

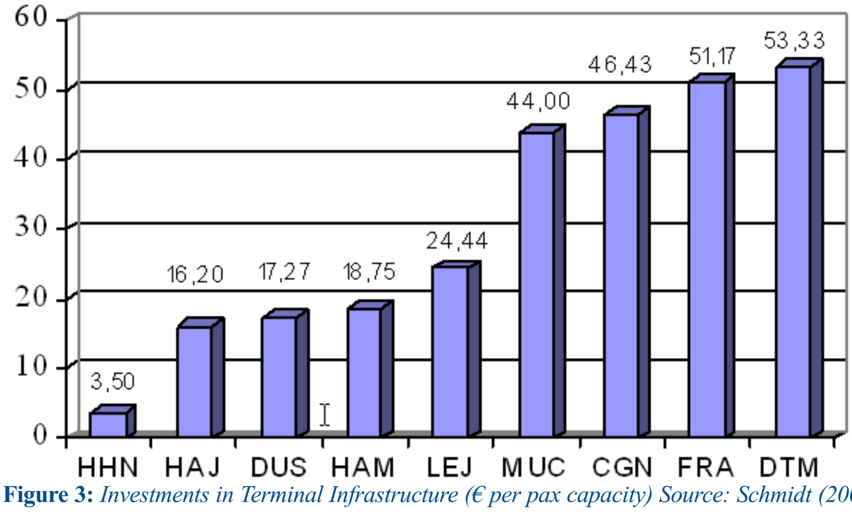


Figure 3: *Investments in Terminal Infrastructure (€ per pax capacity) Source: Schmidt (2005)*

But is DEA an appropriate approach to measure the technical efficiency of airports or do parametric methods such as SFA provide more sufficient results? Firstly, in DEA, one does not have to make any assumptions regarding the distribution of the error term. This is different in SFA as it is a parametric function. Hence, making different assumptions of the error term will automatically lead to different results. Secondly, DEA is a non-parametric approach, where the efficient frontier is constructed by the technically efficient airports, whereas SFA estimates either a cost or production frontier. Furthermore, the question now arises, which of the two is the more adequate method of measuring the technical efficiency of airports? Banker, Gadh and Gorr (1993) discussed this problem in general and came to the result that SFA is more appropriate when severe measurement noise is expected and a cost or production function can be assumed. DEA on the other side seems to be the better approach when measurement errors might not severely affect a firms' performance and the assumption of a neoclassical theory is more doubtful. Nevertheless, "neither method performed satisfactorily for high measurement errors." (Banker et al 1993, p.332). Since the airport's performance might be affected by measurement noise such as crises, weather conditions, specific legal constraints and air traffic management problems, SFA might be the more appropriate method to apply. However, SFA cannot handle multiple outputs when only traffic data is available as the aggregation of passengers, cargo and air transport movements to a single output such as the airport throughput unit (ATU) and also the work load unit is not without its critics. Therefore, including financial data in the sample allows for the consideration of multiple outputs. In conclusion, there is no a priori reason which strikes the balance for one method, it is not clear if, for example, the stakeholders in a regulatory progress will agree on a common benchmarking method. This certainly limits the practical use of benchmarking.

All in all, to receive results that live up to the high demands and expectations of managers and regulators, more work has to be done in the adjustment of the inputs and outputs in the future.

Footnotes

1. The author thanks Christiane Mueller-Rostin and Hans-Martin Niemeier and the other project members of GAP for helpful comments that are gratefully acknowledged. The responsibility for any remaining shortcomings remains the author's.
2. See for example the benchmarking studies by the Air Transport Research Society (ATRS) and Transport Research Laboratory (TRL).
3. The airports that have been included in this analysis are Bremen, Dresden, Dortmund, Duesseldorf, Muenster-Osnabrueck, Frankfurt, Hanover, Hamburg, Leipzig, Munich, Nurem-berg, Saarbruecken, Stuttgart and the Berlin airports Schoenefeld, Tegel and Tempelhof.
4. For more information please visit the projects' website on www.gap-projekt.de.
5. One work load unit equals one passenger or 100kg of cargo.
6. The passenger volume of 2004 is given in brackets.
7. Note that in an output maximizing model the inputs are fixed.

References

Air Transport Research Society (2005), 'Airport Benchmarking Report – 2005: Global Standards for Airport Excellence', Vancouver, ATRS.

Banker, R., Gadh, V. & Gorr, W. (1993), 'A Monte Carlo comparison of two production frontier estimation methods: Corrected ordinary least squares and data envelopment analysis', European journal of operational research, vol. 67, no. 3, pp. 332-343.

Charnes, A., Cooper, W. W. & Rhodes, E. (1978), 'Measuring the efficiency of decision making units', European journal of operational research, vol. 2, no. 6, pp. 429-444.

Coelli, T. J., Prasada Rao, D. S., O'Donnell, C. J. & Battese, G. E. (c 2005), An introduction to efficiency and productivity analysis, Springer: New York, NY.

Farrell, M. J. (1957), 'The measurement of productive efficiency', Royal Statistical Society: Journal of the Royal Statistical Society / A, vol. 120, no. 3, pp. 253-290.

Pels, E., Nijkamp, P. & Rietveld, P. (2001), 'Relative efficiency of European airports', Transport policy, vol. 8, no. 3, pp. 183-192.

Schmidt, L. (2005), 'Fahrt in die Zukunft', fvw spezial, vol.15/05, pp.15-23

Selten, R. (1970), Preispolitik der Mehrproduktunternehmung in der statischen Theorie, Springer: Berlin.

Transport Research Laboratory (2004), 'Airport Performance Indicators', Wokingham, TRL.

Table and Figure Appendix

Table 1: *Technical Efficiency Scores DEA*

	Pax '04 (in millions)		1998	1999	2000	2001	2002	2003	2004	2005
Bremen	1.7	BRE	0.944	0.929	0.892	1.000	1.000	1.000	1.000	0.966
Dresden	1.6	DRS	0.879	0.897	1.000	0.903	0.869	0.870	0.916	0.905
Dortmund	1.2	DTM	1.000	1.000	1.000	0.681	0.681	0.680	0.830	0.839
Duesseldorf	15.2	DUS	1.000	1.000	1.000	1.000	1.000	0.984	1.000	0.998
Muenster-Osnabrueck	1.5	FMO	0.908	1.000	1.000	1.000	1.000	1.000	0.991	0.986
Frankfurt	50.1	FRA	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Hannover	5.2	HAJ	0.732	0.713	0.717	0.691	0.632	0.634	0.650	0.681
Hamburg	9.8	HAM	0.807	0.781	0.771	0.758	0.714	0.850	0.840	0.789
Leipzig	1.9	LEJ	0.936	0.957	0.761	1.000	1.000	0.567	0.570	0.827
Munich	26.7	MUC	0.895	0.921	0.958	1.000	1.000	0.864	0.897	0.934
Nuremberg	3.6	NUE	0.763	0.778	0.822	0.926	0.969	0.961	1.000	0.888
Saarbruecken	0.4	SCN	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Stuttgart	8.7	STR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Berlin-Schoenefeld	3.3	SXF	0.375	0.340	0.363	0.440	0.393	0.427	0.842	0.454
Berlin-Tempelhof	0.4	THF	0.313	0.256	0.224	0.667	0.542	0.367	0.360	0.390
Berlin-Tegel	11.0	TXL	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Mean	**8.9**	∅	**0.938**	**0.942**	**0.934**	**0.947**	**0.941**	**0.917**	**0.948**	**0.938**

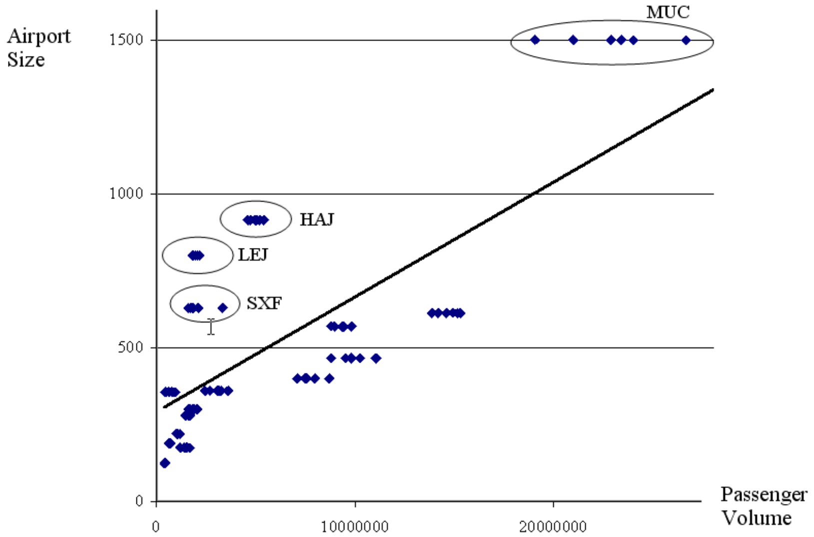


Figure 2: *Passengers per Airport Size (ha)*